Cognition 119 (2011) 313–324

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Seeking positive experiences can produce illusory correlations

Jerker Denrell [b,1,2], Gaël Le Mens [a,*,1]

[a] Universitat Pompeu Fabra, Department of Economics and Business, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain
[b] Saïd Business School, Park End Street, Oxford, Oxfordshire, OX1 1HP, United Kingdom

ARTICLE INFO

Article history:
Received 29 January 2010
Revised 18 January 2011
Accepted 21 January 2011

Keywords:
Reinforcement learning
Sampling
Adaptive behavior
Stereotype formation
Halo effect

ABSTRACT

Individuals tend to select again alternatives about which they have positive impressions and to avoid alternatives about which they have negative impressions. Here we show how this sequential sampling feature of the information acquisition process leads to the emergence of an illusory correlation between estimates of the attributes of multi-attribute alternatives. The sign of the illusory correlation depends on how the decision maker combines estimates in making her sampling decisions. A positive illusory correlation emerges when evaluations are compensatory or disjunctive and a negative illusory correlation can emerge when evaluations are conjunctive. Our theory provides an alternative explanation for illusory correlations that does not rely on biased information processing nor selective attention to different pieces of information. It provides a new perspective on several well-established empirical phenomena such as the 'Halo' effect in personality perception, the relation between proximity and attitudes, and the in-group out-group bias in stereotype formation.

1. Introduction

Much research on illusory correlation has focused on how an individual's processing of information might produce biases in her assessment of the correlation between features or attributes of objects or alternatives. Previous explanations have proposed that prior expectations (Chapman & Chapman, 1967), the differential distinctiveness of positive versus negative stimuli (Allan, 1993) and the greater distinctiveness of infrequent events (Hamilton & Gifford, 1976) distort the encoding and recall of information used to estimate the correlation between features (for a recent review, see Fiedler, 2000a). Some researchers have also proposed that illusory correlation might emerge from skewed distributions of the features in the population (Fiedler, 1991, 2000b; Fiedler, Freytag, & Meiser, 2009) or from the consideration of limited samples (Kareev, 1995a, 1995b, 2000; Kareev & Fiedler, 2006).

Most of these prior approaches assume that individuals have access to information about two or more dimensions and try to explain why the perceived correlation, based on this sample, would diverge from the true correlation. In reality, however, information is not always immediately available. People may have to sample the information sequentially and form beliefs accordingly (Anderson, 1981; Hogarth & Einhorn, 1992; Kashima & Kerekes, 1994; March, 1996). For example, people may only be able to obtain information about the attributes of other individuals by interacting with them. In such contexts, sampling is often adaptive: prior observations usually affect the probability of future sampling (Denrell, 2005; March, 1996; Smith & Collins, 2009). For example, an individual may not want to continue to interact with others unless she believes that the interaction will be productive or pleasurable. In this paper, we show that when decision makers sequentially sample information, the resulting sample bias might produce an illusory correlation even when information is correctly processed.

* Corresponding author. Tel.: +34 93 542 27 17.
E-mail addresses: jerker.denrell@sbs.ox.ac.uk (J. Denrell), gael.le-mens@upf.edu (G. Le Mens).
[1] Both authors contributed equally to this work.
[2] Tel.: +44 1865 28 89 48 (J. Denrell).

0010-0277/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2011.01.007

To explain why this might happen, consider the following example. You are visiting an academic department for a few months and learning about your new colleagues. Suppose that you pay attention to two attributes of your colleagues: their creativity and their agreeableness. You can only learn about these attributes from personal interactions. We assume the probability that you interact with a colleague depends on your estimates of her two attributes. In particular, being mindful of not wasting your time you decide that you will only continue to interact with people who seem to be creative or agreeable.

Each encounter with a colleague provides additional information about her level of creativity and agreeableness. We assume that your estimates of the two attributes of your colleagues are based on your past interactions. More precisely, your estimate of the creativity and agreeableness of a particular colleague is a weighted average of all your past experiences with this individual.

Suppose that creativity and agreeableness are independently distributed in the department. That is, there is no correlation between the attributes – an individual who is creative is not more or less likely to be agreeable than an individual who is not creative. Here we show that you will nevertheless perceive the two attributes to be positively correlated. Specifically, suppose your estimates of the creativity and agreeableness of your colleagues at the end of your visit are $\hat{c}_1, \hat{a}_1, \ldots, \hat{c}_N, \hat{a}_N$. Let $\widehat{C}$ and $\widehat{A}$ be the vectors of estimates of creativity and agreeableness. Under the above assumptions, $\widehat{C}$ and $\widehat{A}$ will be positively correlated.

The key to the emergence of this illusory correlation is that you may stop interacting with a colleague depending on your assessment of her creativity and agreeableness. Suppose you believe colleague i to lack creativity ($\hat{c}_i$ is low). If you find this colleague disagreeable ($\hat{a}_i$ is low), you are unlikely to interact with her again, and thus your belief about her lack of creativity will tend to persist even if she is in fact creative. If, on the contrary, you believe i to be agreeable ($\hat{c}_i$ is low but $\hat{a}_i$ is high), you are likely to interact with her again. Doing so, you might discover that she is in fact creative. Overall, this sequential process of belief formation and information sampling implies that the distribution of estimates at the end of the visit will diverge from the distribution of attributes in the department. Because combinations of estimates that lead to avoidance ($\hat{c}_i$ low, $\hat{a}_i$ low) will be more stable than combinations of estimates that lead to further sampling (e.g. $\hat{c}_i$ low, $\hat{a}_i$ high), combinations of estimates that lead to avoidance will be over-represented.

In this example, the illusory correlation was positive. But, had we assumed you only wanted to interact with people whom you believed to be *both* creative and agreeable rather than just creative *or* agreeable, the illusory correlation would have been negative. In this case, only high estimates regarding both creativity and agreeableness ($\hat{c}_i$ high, $\hat{a}_i$ high) lead to further sampling. Because estimates that lead to further sampling are less stable than estimates that lead to avoidance, the combination ($\hat{c}_i$ high, $\hat{a}_i$ high) will be under-represented in the distribution of final estimates relative to the distribution of attributes in the department. It follows that un-balanced combinations ($\hat{c}_i$ high, $\hat{a}_i$ low or $\hat{c}_i$ low, $\hat{a}_i$ high) will be over-represented in the distribution of estimates at the end of the visit relative to balanced combinations ($\hat{c}_i$ high, $\hat{a}_i$ high or $\hat{c}_i$ low, $\hat{a}_i$ low). Such an asymmetry corresponds to a negative correlation.

More generally, the sign of the illusory correlation depends on how the decision maker combines estimates in making her sampling decisions. In the rest of the paper, we develop and analyze a simple learning model to clarify the underlying mechanism. In the next section, we first describe our model and show that the sign of the illusory correlation depends on the joint influence of attribute estimates on the sampling rule. Then we further clarify the role of the sampling rule. We show that a positive illusory correlation emerges when evaluations are compensatory (an object is evaluated positively if the average of the two features is positive) or disjunctive (an object is evaluated positively if one of two features is positive) and that a negative illusory correlation can emerge when the evaluation of an object is conjunctive (an object is evaluated positively only if both features is positive). We then discuss how our mechanism differs from prior explanations that rely on biased information processing and from prior sampling-based explanations. We also discuss the scope of our model and what happens when some of the assumptions are relaxed. In particular, we show that our model also applies to settings when the true correlation between attributes is distinct from zero. In such cases, the form of the sampling rule leads to a tendency to overestimate or underestimate the true correlation. Finally, we explain how our model helps cast a new light on well-known psychological phenomena such as the 'Halo' effect in person perception, the formation of stereotypes and the well-documented empirical finding that people tend to like proximate others more than distant others.

2. Model

To illustrate how illusory correlations can emerge as a result of adaptive sampling, we develop a model in which an individual learns about the values of two attributes of an alternative from her experiences with that alternative (see Fig. 1). Consistent with the idea of adaptive sampling, we assume the probability of selecting the alternative is a function of the decision maker's estimates of the values of the two attributes, denoted by $\hat{x}_t$ and $\hat{y}_t$. Let $Q(\hat{x}_t, \hat{y}_t)$ denote this probability of selecting the alternative given the estimates. Our model is based on the following assumptions:

(a) In each period, the decision maker can select the alternative or not. If the decision maker selects the alternative, she observes both attributes in that period. If the decision maker does not select the alternative, she does not get any information about the values of the attributes.
(b) The observations of the two attributes are *independent* realizations of two *independent* random variables with positive variances. In particular, past observations of an attribute do not affect how new information about that attribute is interpreted and encoded. In addition, observations of attribute 1 (resp. attribute 2) do not affect how information about attribute 2 (resp. attribute 1) is interpreted.

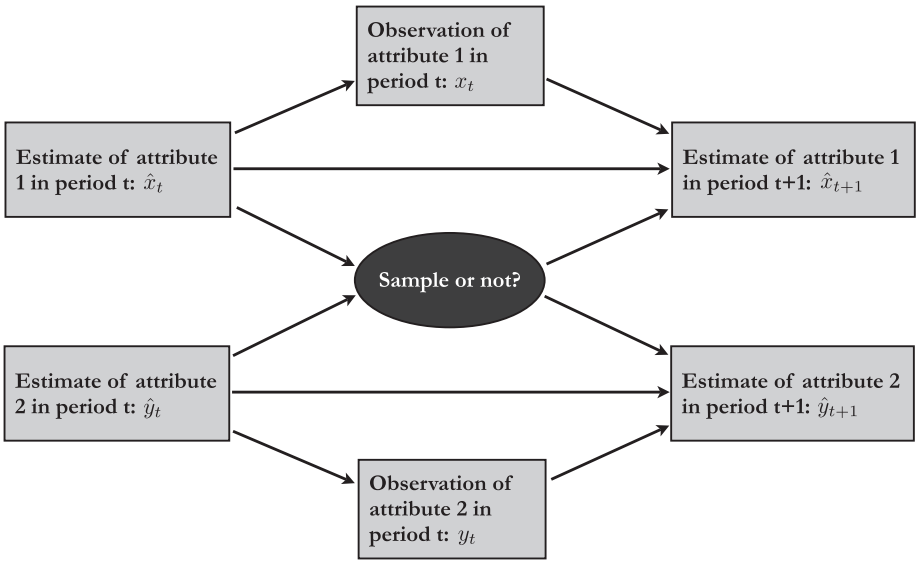

Fig. 1. Structure of the reinforcement learning process: If the alternative is selected, the decision maker updates her estimate of the value of an attribute on the basis of her observation of that attribute. The estimates do not change if the alternative is not selected. The sampling decision depends on both estimates.

(c) The estimate of the value of an attribute at the beginning of period t is any weighted average of the past observations of that attribute. Specifically the revised estimate of the first attribute after an observation has been made in period t can be written $\hat{x}_{t+1} = (1 - b_t)\hat{x}_t + b_t x_t$, where x_t denotes the observation of the first attribute in period t. The weight of new information is b_t, with $0 < b_t < 1$. The same updating rule applies to the estimate of the second attribute. The initial estimates are random draws from the underlying distributions of observation values. Estimates do not change when no observation is made. Denrell (2005) has shown this weighted average model provides a reasonable approximation to actual belief updating in several experimental studies. In addition, the model can be derived from more realistic connectionist models (e.g. Busemeyer & Myung, 1992).

(d) The probability of sampling the alternative in period t, $Q(\hat{x}_t,\ \hat{y}_t)$, is twice continuously differentiable [3] and is an increasing function of both arguments. The idea is that the decision maker values the attributes positively.

In summary, the decision maker learns about the values of the two attributes on the basis of her observations of these attributes and the decision maker is more likely to sample an alternative if she has a positive estimates of its attributes. An important assumption of the model is that when the decision maker selects the alternative, her observations of the two attributes are *uncorrelated*. In other words, the underlying correlation between the two attributes is equal to zero.

The above model leads to the emergence of an illusory correlation: the estimates $\hat{x}_t$ and $\hat{y}_t$ become correlated. The sign of the illusory correlation depends on how the attributes interact in driving the sampling probability, as formulated in the following theorem. (i) states that if there is a negative interaction between the attributes (i.e., the effect of an attribute on the sampling probability decreases when the other attribute becomes more positive), then the illusory correlation will be positive; (ii) states that if there is a positive interaction between the attributes, then the illusory correlation will be negative; and (iii) states that if the two attributes have independent effects on the sampling probability, no illusory correlation will emerge.

Theorem 1. *Suppose conditions (a)–(d) hold.*

(i) *If* $\frac{\partial^2}{\partial \hat{x}_t \partial \hat{y}_t} \log Q(\hat{x}_t, \hat{y}_t) < 0$ *for all* $\hat{x}_t,\ \hat{y}_t \in \mathbb{R}$, *then* $\hat{x}_t$ *and* $\hat{y}_t$ *become positively correlated for t large enough.*

(ii) *If* $\frac{\partial^2}{\partial \hat{x}_t \partial \hat{y}_t} \log Q(\hat{x}_t, \hat{y}_t) > 0$ *for all* $\hat{x}_t,\ \hat{y}_t \in \mathbb{R}$, *then* $\hat{x}_t$ *and* $\hat{y}_t$ *become negatively correlated for t large enough.*

(iii) *If* $\frac{\partial^2}{\partial \hat{x}_t \partial \hat{y}_t} Q(\hat{x}_t, \hat{y}_t) = 0$ *for all* $\hat{x}_t,\ \hat{y}_t \in \mathbb{R}$, *then the correlation between* $\hat{x}_t$ *and* $\hat{y}_t$ *becomes arbitrarily close to zero for t large enough.*

Proof. This formal result has been proven by Denrell and Le Mens (2007) in another context.[4] It relies on the fact that it is possible to write an explicit formula for the asymptotic distribution of estimates of the values of the attributes. While it is impossible to derive a simple formula to express the probability distribution of the estimates in any given period, computer simulations confirm that an illusory correlation emerges after a reasonably low number of periods (see Section 3). □

[3] This regularity condition is necessary to keep the formalism simple. Similar results hold even if this condition is not satisfied (see Karlin & Rinott (1980) for underlying mathematical results).

[4] Similar results hold when there are more than two attributes. The extension is straightforward for the positive illusory correlation case, but much more complicated for the negative illusory correlation case (see Karlin & Rinott (1980), for a detailed discussion of the underlying mathematical formalism).

To understand the intuition underlying this result note first that the sampling process implies that each of the attributes tends to be underestimated. The reason is that a negative estimate reduces the probability of sampling, which implies that no further information is available and the estimate will not be updated (Denrell, 2005; Fazio, Eiser, & Shook, 2004; March, 1996; Smith & Collins, 2009).

In the above model, however, the probability that an alternative is sampled depends on both attributes. This raises the possibility that an alternative continues to be sampled, even if one of the attributes is perceived to be low because the other is perceived to be high. In this case, the estimate of the first attribute will be updated and is likely to increase (due to regression to the mean). Thus, a high estimated value of the second attribute has a systematic positive effect on the estimate of first attribute.

If the estimate of the first attribute is initially high, the effect of the second attribute works in the opposite direction: a high estimate of the value of the second attribute increases the probability of sampling and, due to regression to the mean, the first estimate is likely to decrease. The above theorem shows that the relative magnitudes of these two effects depends on the precise way in which the decision maker combines the estimates of the two attributes in her decision to select the alternative. This, in turn, determines the sign of the illusory correlation.

3. Positive or negative illusory correlation: the role of the sampling rule

In this section, we provide a more explicit interpretation of the conditions that determine the sign of the illusory correlation by analyzing more specific sampling rules. We assume that the decision maker combines her estimates of the two attributes in a single estimate of the overall value of the alternative. This overall subjective value, or 'utility' (as generally referred to in the literature on multi-attribute decision making, e.g. Keeney & Raiffa, 1993) is then used as the basis for the sampling decision. More precisely, we will assume that an alternative with a higher utility is more likely to be selected than an alternative with a lower utility. The analysis focuses on how the form of the multi-attribute utility function impacts the sign of the illusory correlation.

In the example of the introduction we assumed you wanted to continue to interact with colleagues you perceived to be creative *or* agreeable. In other words, you saw these two attributes as *substitutes* to each others. Another possibility is that you only want to interact with people that you find *both* creative and agreeable. We now specify the probability of sampling, $Q(\hat{x}_t, \hat{y}_t)$, to take into account such considerations about how the decision maker combines attributes. Specifically:

(e) The decision makers' two-attribute 'utility' function is defined by $\hat{u}_t = U_p(\hat{x}_t, \hat{y}_t) = [0.5(\hat{x}_t)^p + 0.5(\hat{y}_t)^p]^{1/p}$ where $p \in \mathbb{R}$, $0 < \hat{x}_t < \alpha$ and $0 < \hat{y}_t < \alpha$, for some $\alpha > 0$ (i.e. the estimates of the attributes are assumed to be positive and bounded random variables). This utility function, the 'root power mean' or the 'quasi-linear' mean has been used in previous work (Dawes, 1964; Einhorn, 1970) to represent both conjunctive, compensatory, and disjunctive utility functions. It has the following properties (see also Fig. 2):
- For all $p \in \mathbb{R}$, $\min(x, y) \leqslant U_p(x, y) \leqslant \max(x, y)$.
- If $p > 1$, it puts more weight on the maximum of x and y. Thus, when $p > 1$ the utility function is 'compensatory': the two attributes tend to substitute for each other. Moreover, it can be shown that $\lim_{p \to +\infty} U_p(x,y) = \max(x,y)$. This corresponds to a situation where the utility function is disjunctive.
- If $p = 1$, then it is the arithmetic average: $U_1(x,y) = 0.5x + 0.5y$.
- If $p < 1$, then it puts more weight on the minimum of x and y. Thus, when $p < 1$, the utility function is 'non-compensatory'. Moreover, it can be shown that $\lim_{p \to -\infty} U_p(x,y) = \min(x,y)$. This corresponds to a situation where the utility function is conjunctive.

(f) The probability of sampling depends on the 'utility' through the logistic choice rule: $Q(\hat{x}_t, \hat{y}_t) = 1/(1 + \exp[-s(\hat{u}_t - \gamma)])$, where s is a parameter that regulates the sensitivity of the choice probability to the utility, and γ is a scaling parameter. The logistic choice rule has often been used to model choices under uncertainty (e.g. Luce, 1959).

Given these assumptions, when is there a positive or negative correlation? The following proposition provides an answer to this question.

Proposition 2. *Suppose conditions (a)–(f) hold. The correlation between $\hat{x}_t$ and $\hat{y}_t$ is positive for t large enough whenever $p \geqslant 1$. Moreover, there is a value of $p^* < 1$ such that the correlation becomes negative for $p < p^*$ and for t large enough.*

Proof. See the Appendix. □

Thus, whether the illusory correlation is positive or negative depends on how the two attributes are combined in the overall valuation of the alternative. If the attributes are substitutes $\left(\frac{\partial^2 \hat{u}_t}{\partial \hat{x}_t \partial \hat{y}_t} \leqslant 0,\ p > 1\right)$, a positive illusory correlation emerges. When the attributes are *strong* complements (p is low enough), a negative illusory correlation emerges. The assumption of complementarity, $\left(\frac{\partial^2 \hat{u}_t}{\partial \hat{x}_t \partial \hat{y}_t} > 0\right)$, is by itself not enough to guarantee the emergence of a negative illusory correlation.

The picture is clearer when one consider the extreme cases of disjunctive and conjunctive sampling rules. When the rule is disjunctive, ($\hat{u}_t = \max(\hat{x}_t, \hat{y}_t)$), a positive illusory correlation emerges. When the rule is conjunctive ($\hat{u}_t = \min(\hat{x}_t, \hat{y}_t)$), a negative illusory correlation emerges. Finally, note that in the special case where the value of the alternative is the arithmetic average of the attribute values ($p = 1$), there is a positive illusory correlation (this is a special case of compensatory utility function).

The above theoretical results only hold asymptotically, but simulations show that illusory correlations tend to emerge relatively quickly. Fig. 3 displays the size of the illusory correlation after 10 and 50 periods, as a function of the parameter p. The amplitude of the illusory correla-

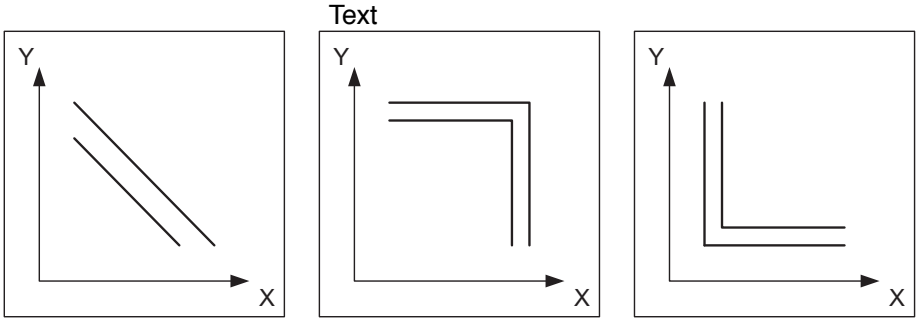

Fig. 2. Iso-utility curves for quasi-linear mean utility. The left quadrant corresponds to $p = 1$, the middle quadrant to $p = +\infty$ and the right quadrant to $p = -\infty$.

tion can be substantial, even after only 10 periods. As discussed above, when p is equal to or higher than 1, the correlation is positive. Also, the higher p (the more compensatory the utility function), the higher the correlation. Similarly, the correlation is negative for low values of p.

4. Relation to existing research on illusory correlation

As noted in a recent review (Fiedler, 2000a), research on the misperception of correlations has evolved in separate areas of psychology such as personality research, stereotype formation, associative learning or applied psychology. For the purpose of this paper, we classify prior explanations of illusory correlations in two categories: the theories that explain systematic errors in perception of correlations by invoking a bias in how the mind processes available information, and sample-based explanations. We discuss, in turns, how our explanation differs from these two classes of prior explanations.

4.1. Theories based on biased information processing

As mentioned in the introduction, most prior explanations of illusory correlations rely on a bias in how the information about the relation between variables is processed. In the standard paradigm, experimental participants observe a set of items each characterized by a pair of attribute values (X, Y).[5] Information available to participants can be summarized in a 2×2 contingency table such as that of Fig. 4. x_1 and x_2 are the two possible values for X and y_1 and y_2 the two possible values for Y. The sign of the correlation between the two variables depends on the products of the cells in the two diagonals. When $ad > bc$, the correlation is positive (x_1 is likely to co-occur with y_1 and x_2 is likely to co-occur with y_2); when $ad < bc$, the correlation is negative (x_1 is likely to co-occur with y_2 and x_2 is likely to co-occur with y_1); and when $ad = bc$, the correlation is equal to zero (the two variables are independent). After having observed the $N = a + b + c + d$ items, participants are asked about the contingency, or correlation, between the two features. The theories that rely on information processing biases explain illusory correlation by proposing that some observations receive more weight in the computation of the correlation

[5] Many studies consider one of the variables to be a predictor, and the other to be a criterion. Since this distinction is irrelevant for the foregoing discussion, we refer to the two variables as feature values.

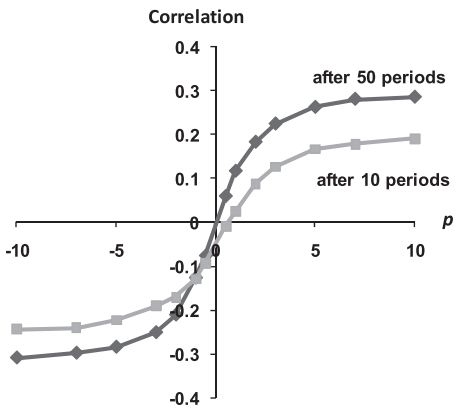

Fig. 3. Correlation between $\hat{x}_t$ and $\hat{y}_t$ as a function of the shape of the choice rule (as characterized by p). Based on 10,000 computer simulations of a version of the learning model with the quasi-linear mean utility function. The observations are independent realizations of two random variables with uniform distributions between 0 and 1. Figure obtained with $b_t = 0.5$, $s = 10$, and $\gamma = 0.5$.

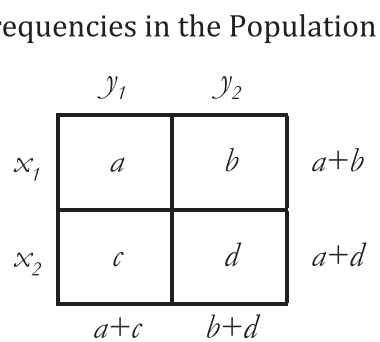

Fig. 4. Standard contingency table in studies on illusory correlation. The two variables X and Y are binary. X can take value x_1 or x_2 and Y can take value y_1 or y_2. Each item is represented by a pair (X, Y). The values a, b, c and d represent the frequencies of each combinations of attribute values. The population is made of $N = a + b + c + d$ items.

than others. For example, if items in cell a receive a disproportionate weight, such as when these items are rare and distinctive (e.g. Hamilton & Gifford, 1976), negative and distinctive (e.g. Allan, 1993) or consistent with expectations (e.g. Chapman & Chapman, 1967) the perceived correlation is likely to be positive.

To adapt our model to this 2×2 design, suppose that X and Y refer to creativity and agreeableness as in the introductory example and there is a threshold, such that if

$X \geqslant \theta_X$ an individual is considered creative (x_1), and if $X < \theta_X$ the individual is considered uncreative (x_2). Similarly, there is a threshold for the assignment of individuals as agreeable (y_1) or disagreeable (y_2).

There are two important differences between our model and prior explanations. First, we do not assume that the decision-maker has a direct access to attribute values of the N items. Rather, the decision-maker can only observe noisy signals of the attribute values. For example, the decision maker can only observe if an individual she meets was creative and agreeable during that interaction. This implies that the decision-maker's estimates of the attributes of this individual are subject to noise. If the decision maker continues to interact with the same individual, however, the estimates are likely to become more accurate.

Second, in addition to assuming noisy observations of the attribute values, we assume that the decision-maker learns about the attribute values through a process of sequential sampling. That is, after each interaction with an individual (or object) the decision-maker decides whether to continue sampling. Moreover, the sampling probability depends on the estimated attribute values. The decision-maker is more likely to interact again with an individual perceived to be creative than with an individual perceived to be uncreative. Similarly, the decision-maker is more likely to interact again with an individual perceived to be agreeable than with an individual perceived to be disagreeable.

These two assumptions imply that the distribution of estimated attribute values will diverge from that in the population. To illustrate, suppose the decision maker evaluates 20 individuals. Assume that people are in fact equally likely to be creative or not and equally likely to be agreeable or not and that the true correlation between these two attributes is zero. The decision maker does not observe the creativity or agreeableness of the 20 individuals directly, but can only learn about them when she meets these 20 individuals. However, because observations are noisy, individuals may be misclassified: based on the first interaction, the decision maker may mistakenly believe, for example, that someone is not creative when that individual is in fact creative. It is also possible, of course, that someone who is not creative might be classified as creative. But this misclassification will often be found out and corrected because it leads to additional sampling. The mistake of classifying someone as uncreative will seldom be corrected, however, because the decision-maker tends to avoid individuals classified as uncreative.

More generally, the avoidance of alternatives with attributes believed to have low values leads to an overestimation of the prevalence of attribute values lower than the true average. This implies that the majority of the 20 individuals will be classified as uncreative. Similarly the majority of the 20 individuals will be classified as disagreeable. In other words, the marginal distribution of the estimates of the attribute values will be skewed despite the fact that the marginal distributions of the attribute values in the population are symmetric.

How about the joint distribution of attribute estimates in the population? Our theorem predicts different patterns depending on the form of the sampling rule. If, for example, the sampling rule satisfies condition (i), the joint distribution of estimates will imply a positive correlation. For example, the decision-maker might come to believe that there are five individuals that are both creative and agreeable, four that are creative but disagreeable, four that are uncreative but agreeable, and seven that are neither creative nor agreeable (e.g. Fig. 5, left panel). But if the sampling rule meets condition (ii) the estimates will be negatively correlated (e.g. Fig. 5, right panel). Finally, if the sampling rule meets condition (iii) the estimates will be independent.

It should be clear from this discussion that our model does not assume that items in some cell(s) receive more or less weight in the computation of the correlation estimate. Rather, the illusory correlation emerges because of differences in propensities to mis-categorize some items, due to the assumption of adaptive sampling. This, in turn, leads to a correlation between estimates that is distinct from the true correlation. The pattern of mis-categorization and, thus, of illusory correlation, depends on the form of the sampling rule. In summary, the crucial difference between our mechanism and explanations that rely on differential weighing of some observations is that it places the locus of the bias *outside* the head of the decision maker. As such, it does not contradict theories that rely on biased information processing. Rather, it suggests that in environments where decision-makers have to learn about the attribute values from their own experiences, information processing biases might operate on top of sampling biases. Whether sampling biases or information processing biases are more important in explaining illusory correlation in real-world environments is an open question.

This discussion suggests that our explanation applies primarily to illusory correlation between distal entities (constructs that are not directly observable, such as personality traits, intelligence, danger, similarity and health). By contrast, our model cannot explain illusory correlation between proximal entities (constructs that are directly observable). The reason is that, in this later setting, there is almost no possibility for errors in estimates (and thus of categorization). But, importantly, our model can also apply to settings where one of the construct is distal and other one is proximal. This setting is of particular relevance to research on stereotype formation and maintenance (e.g. Hamilton & Gifford, 1976). We discuss this topic in the section on Stereotype Formation.

4.2. Relations to other sampling explanations

Several investigators have also proposed that illusory correlation might originate from a bias in the sample of information individuals have access to. Here, we discuss the similarities and differences between these prior contributions and our approach.

4.2.1. The small-sample effect

In a series of papers, Kareev (1995a, 1995b, 2000) have demonstrated that correlations estimated from small samples tend to be larger than correlations in the population. The main reason is that the distribution of correlation estimates tends to be skewed in such a way that most

Positive Correlation

	y_1	y_2	
x_1	5	4	9
x_2	4	7	11
	9	11	

Negative Correlation

	y_1	y_2	
x_1	3	6	9
x_2	6	5	11
	9	11	

Fig. 5. Examples of possible categorical assignments implied by the adaptive sampling model. Marginal distributions are skewed and the same for the two tables. The pattern of joint frequencies depend on the choice rule, as formalized in the theorem. The positive correlation value is 0.19 and the negative correlation value is −0.21.

estimates are above the true correlation (when the true correlation is large). Kareev proposes this is important because correlations are likely to be estimated from limited samples, due to the limited capacity of the working memory. To understand the difference between Kareev's analysis and ours, it is useful to note that in Kareev's papers, 'small sample' refers to *N*, the number of items considered for the computation of the correlation. In our model, however, the key is that information samples about the *N* items can be biased in systematic ways. Our model suggests that an illusory correlation can emerge even if the decision maker uses a large sample, in the sense that it contains many items.

4.2.2. Non-proportional sampling

Kareev and Fiedler (2006, see also Fiedler, 2000b) have examined the influence of sampling behavior on correlation estimates. They considered the relation between a skewed binary predictor and a binary criterion variable. They showed that a nonproportional sampling scheme (a sample that has a marginal distribution of the predictor less extreme than the marginal distribution in the population) leads to a tendency to overestimate the correlation between predictor and criterion. This explanation is similar to ours, in a sense, because it assumes that people have access to a biased frequency table. The source of the sampling bias is different, however. Kareev and Fiedler analyzed a situation where people have an information-search goal, whereas we analyze a situation where the decision maker has a hedonistic goal: her sampling strategy is motivated by the desire to avoid negative experiences, not by the goal to select the most informative alternative.

4.2.3. The effect of environmental base rates

Several papers by Fiedler and colleagues have analyzed the influence of skewed base rates on the perception of correlation and contingency (Fiedler, Walther, & Nickel, 1999; Fiedler, Walther, Freytag, & Plessner, 2002; Fiedler & Freytag, 2004; Fiedler, Freytag, & Unkelbach, 2007). These culminate in a theory of pseudocontingencies (PC) which explains why people tend to believe that variables with skewed marginal distributions tend to covary (i.e. have a positive correlation) even when they do not make joint observations of the pairs of variables (Fiedler et al., 2009). Fiedler and colleagues explain PC inferences as mistaken contingency statements made at the individual level on the basis of an ecological correlation that holds at a more aggregated level. As such, it is the result of a cognitive illusion that does not pertain to differential weighing of some pieces of information, but rather to a misunderstanding of the mechanism that generated the available sample of information. This leads to the consideration of irrelevant information in the computation of the contingency or correlation. PC inferences occur when base rates (i.e. marginal distributions) are skewed (Fiedler et al., 2007). In our model, adaptive sampling leads to skewed marginal distributions. This suggests that adaptive sampling as described here might provide a context for PC inferences: when people learn from experience and can avoid alternatives that lead to poor outcomes, PC inferences likely operate. How PC inferences combine with the effect of adaptive sampling is, however, an open question.

The main issue is that several joint distributions are consistent with a given pair of marginal distributions, as illustrated by Fig. 5. In our analyses, we assumed that people were able to correctly compute the correlation between their attribute estimates. Fiedler and colleagues' investigations of PC inferences, however, suggest that this computational step is prone to systematic errors. More precisely, the PC algorithm suggests that people might have a tendency to see a positive correlation even in situations where the correlation between their estimates is null or even negative, as in the table of the right panel of Fig. 5. To what extent this actually happens deserves empirical inquiry.

5. Discussion of the assumptions of the model

While the model developed in this paper makes only minimal assumptions about how information is stored in memory, it relies on important assumptions regarding access to information from the environment. We now discuss these.

5.1. Reinforcement learning structure of the model

We assumed, consistent with the reinforcement learning tradition, that access to information is contingent on actively selecting an alternative and that selection is based

on the valence of past experiences. But access to information does not always satisfy these assumptions. Decision-makers can sometimes observe an alternative without actively selecting it. For example, a decision-maker might obtain information about other people by discussing with a common acquaintance. Also, the decision to select an alternative might be influenced by factors other than the valence of the estimates of its attributes. For example the decision-maker might want to develop accurate estimates of the attributes. When this happens, she will tend to select objects she is more uncertain about, and this sampling decision might be unrelated to the valence of the estimates. There are many determinants of sampling decisions such as curiosity (Loewenstein, 1994), salience (Nisbett & Ross, 1980) and availability (Tversky & Kahneman, 1973). These other factors tend to reduce the adaptive character of the information acquisition process. It is possible to represent such influence by assuming that the probability of sampling in a given period is

$$r + (1 - r)Q(\hat{x}_t, \hat{y}_t).$$

As discussed by Denrell (2005), r can be seen as the probability of nonvoluntary exposure (Thibaut & Kelley, 1959). Even such exposures can carry valuable information about the alternative. If $r = .2$ and $\hat{u}_t = \max(\hat{x}_t, \hat{y}_t)$, the illusory correlation between estimates after 10 periods is about .15, which is lower than the illusory correlation (.21) when there is no such exogenous influence. And the higher the exogenous influence, the lower the illusory correlation. For example, the illusory correlation after 10 periods is only equal to .04 when $r = .7$. More generally, the size of the illusory correlation is increasing in the sensitivity of the probability of making a new observation to past experiences. In the context of reinforcement-learning situations, this pertains to how the decision maker handles the exploration-exploitation trade-off (Sutton & Barto, 1998): more explorative policies will be less conducive to the emergence of illusory correlation, but more exploitive policies will lead to stronger illusory correlation.

We also assumed that the sampling probability is increasing in both arguments (assumption (d) above). We made this assumption to facilitate the exposition of the intuition underlying our main result. But it is in fact not necessary for the illusory correlation to emerge. The only thing that matters for the sign of illusory correlation is whether the joint estimate distribution is log-submodular or log-supermodular (Karlin & Rinott, 1980) and this is uniquely determined by the cross derivative of the logarithm of the sampling probability. This suggests that our model can also apply to settings where the probability of sampling is decreasing in the estimates of the attributes such as when negative traits are valued by the individual forming the estimates. This can happen, for example, when a journalist seeks negative information about a politician to uncover his or her wrongdoings (Peeters, 1983).

5.2. Correlation between observations

We assumed that the true correlation between attribute values is zero to make the proof of the theorem tractable. But computer simulations show that a similar result holds when the correlation between actual attribute values is not equal to zero. In this case, people will tend to *underestimate* or *overestimate* the correlation, depending on the form of the choice rule. Consider what happens when the observations of the two attributes follow bivariate normal distributions with means 0, standard deviations 1, and a correlation of 0.2. If $\hat{u}_t = \max(\hat{x}_t, \hat{y}_t)$, the correlation between estimates is about 0.37 after 10 periods (based on 10,000 simulations and with $s = 10$, $\gamma = 0$ and $b_t = .5$). [6] If $\hat{u}_t = \min(\hat{x}_t, \hat{y}_t)$, then the correlation between estimates is negative rather than positive, at about −.14 after 10 periods. These results suggest that adaptive sampling will generally lead to a misperception of correlations, not only to the detection of a correlation where there is none.

5.3. Number of attributes

We assumed that the alternatives were characterized by two attributes, but similar results hold, to a certain extent, when there are more attributes. Consider, for example, an extension of the quasi-linear mean utility function to N dimensions ($N > 2$). In that case $\hat{u}_t = \left[\frac{1}{N}\sum_{i=1}^{N}(\hat{u}_t^i)^p\right]^{1/p}$. It is possible to show that if $p \geqslant 1$, a positive illusory correlation emerges between all pairs of estimates. It is much more complicated to analyze formally the emergence of negative correlations for three or more attributes.[7]

6. Additional implications

6.1. Halo effect

A vast amount of research has documented a 'halo' effect when individuals rate others on several attributes: the correlation between estimates of positively valued attributes tends to be higher than what is warranted by the data (e.g. Cooper, 1981; Landy & Sigall, 1974; Newcomb, 1931; Nisbett & Wilson, 1977; Thorndike, 1920).

It has been proposed that the halo effect is due to a 'pollution' of attribute ratings by general assessments or assessments of other attributes. This can be due to the fact that people hold implicit personality theories that "nice people tend to have nice attributes and less nice people have less nice attributes" (Nisbett & Wilson, 1977). It has also been proposed that general evaluations affect the interpretation of information about the individual attributes, even when this information should not be ambiguous (Nisbett & Wilson, 1977).

As should be clear from the example of the introduction, about evaluations of creativity and agreeableness, our model helps explain the halo affect. Under the (reasonable) assumption that positive attribute estimates affect the probability of future interpersonal interactions in a mostly compensatory way, people will come to believe that those who have a positive attribute are more likely

[6] Unless otherwise stated, the numerical results that follow are based on simulations using these assumptions.

[7] To understand the complications, note that some patterns of pairwise negative correlations are not possible: for example all pairwise correlations cannot equal −1 when there are more than two attributes.

to also possess another positive attribute. More generally, our model helps explain why positive traits tend to 'go together' in people's personality theories.

How does this relate exactly to the numerous experiments on the halo effect? At first sight, our model does not seem to have much to say about the results of experiments, because the investigator generally controls what information participants have access to, whereas our model relies on the fact that participants choose alternatives based on their own estimates. But most experiments on the halo effect in fact just reveal participants' existing personality theories. For example, a finding that participants judge more favorably an essay written by an attractive person (e.g. Landy & Sigall, 1974) is the signal that attractiveness and competence are associated attributes in the participants' personality theories. And because our model helps explain the formation of haloed personality theories, it provides a foundation for the experimental findings about the halo effect in the evaluation of personality traits.

6.2. Proximity and attitudes

Numerous studies in psychology and sociology show that people develop more positive opinions of proximate others. For example, college students tend to have more positive opinions of their roommates than of other students (Festinger, Schachter, & Back, 1950) even if roommates have been randomly assigned (Marmaros & Sacerdote, 2006; Segal, 1974). Studies also show that managers tend to have more positive opinions of those who work in close proximity to them (Ferris, Judge, Rowland, & Fitzgibbons, 1994) and are also less likely to dismiss those (Landier, Nair, & Wulf, 2007).

Building on the well-accepted expectancy–value model of attitudes, we see attitudes as summary evaluations on an attribute dimension such as good-bad, or likable-dislikable (for a review, see Ajzen, 2001). The above findings can thus be interpreted as instances of illusory correlation between proximity and assessments on the good-bad attribute dimension. Our model predicts that a positive correlation between proximity and attitudes will emerge when people are more likely to interact with others that they like or that are proximate (by Proposition 2). This type of compensatory sampling rule is reasonable, because people will often tend to seek interactions with others they like, even if they are relatively distant. But even if they try to avoid those who are proximate and that they do not like, they might not be able to do so. For example, it is difficult to avoid interacting with one's roommate, even if one does not like him or her.

6.3. Stereotype formation

The model can also explain why individuals might come to associate group membership – which can be viewed as an attribute – with positive or negative characteristics. Suppose, for example, that a decision maker learns, from her experience, about a positively valued attribute, such as creativity, of members of two groups A and B. Let $\hat{c}_t^i$ be the decision maker's estimate of the creativity of individual i at the beginning of period t. Suppose the distribution of actual creativity levels is the same in the two groups and the initial estimate of the creativity of individual i is a random draw from the underlying creativity distribution.

Suppose, however, that the sensitivities of the sampling probability to the creativity estimate differ according to group membership. That is, suppose that $\frac{\partial Q_A(\hat{c}_t^i)}{\partial \hat{c}_t^i} < \frac{\partial Q_B(\hat{c}_t^i)}{\partial \hat{c}_t^i}$ where $Q_A(\hat{c}_t^i)$ (resp. $Q_B(\hat{c}_t^i)$) is the probability that the decision-maker interacts with i in period t if i belongs to group A (resp. B). Such difference in the sensitivity of the probability of interaction with respect to the creativity estimate (or more generally the estimate of an attitude-relevant attribute) can for example occur when A is an in-group and B an out-group (see Denrell (2005), for a review). The reason is that an individual might have to interact with others in the in-group (such as family members or colleagues) even if she finds them uncreative. But because interactions with members of the out-group more directly depend on her decisions, she will tend to interact significantly more with members of the out-group that she finds creative than with members of the out-group that she finds uncreative.

The probability that the decision maker 'samples' another individual thus jointly depends on group membership and the current estimate of creativity. By design, group membership and creativity are independent and thus uncorrelated. Despite this, an illusory correlation will emerge: Estimates of creativity will be more favorable for the members of group A than for the members of group B.[8] To further explain the intuition underlying this result, it is useful to consider a specific case. Suppose that there are a total of 20 individuals, with 5 creative individuals and 5 uncreative individuals in each group. The true correlation between group membership and creativity is zero. Suppose the decision-maker has to interact with members of group A whatever her creativity estimate $\left(\frac{\partial Q_A(\hat{c}_t^i)}{\partial \hat{c}_t^i} = 0\right)$ but she is likely to avoid members of group B she finds uncreative $\left(\frac{\partial Q_B(\hat{c}_t^i)}{\partial \hat{c}_t^i} > 0\right)$. The distribution of creativity estimates for members of group A will be close to the true one because the decision-maker will ultimately learn about the creativity of the members of this group. But because the decision-maker avoids uncreative members of group B, she will likely end up believing that most members of group B are uncreative. For example, she might believe that six members of B are uncreative whereas four members of that group are creative. In terms of the generic contingency table of Fig. 4, we have $a = 5$, $b = 5$, $c = 6$, $d = 4$. This table implies that group membership and creativity assessments are correlated.

This pattern is consistent with the empirical findings that members of an ethnic group tend to have more positive opinions about members of their own groups than of members of other groups (Hewstone, Rubin, & Willis, 2002; Levin, van Laar, & Sidanius, 2003). It is worth noting that, in this setting, saying that the correlation is positive

[8] The proof relies on an extension of Theorem 1 to settings where the sampling probability is not continuous in the attribute values (see Denrell & Le Mens (2007), Appendix B). In this case, the continuity assumption is not satisfied because one of the attributes, group membership, takes discrete values.

or negative is meaningless because the groups do not necessarily have distinct valences. All what matters is that the *sensitivity* of the sampling probability to the distal attribute ('creativity' in the above example) consistently differs across groups.

This example illustrates how some stereotypes, seen as an illusory correlation between group membership and a positively (or negatively) valued attribute, can emerge from experiential learning and adaptive sampling. Contrary to the theories that rely on the distinctiveness of some observations (e.g. Hamilton & Gifford, 1976), our mechanism does not have to assume that marginal distributions (of group membership and of the other relevant attribute) are skewed. It accomodates both settings where the marginals are skewed and settings where marginals are not skewed.

If the sampling bias is at the source of mistaken stereotypes about the out-group, this has important normative implications. In particular, it suggests that, to overcome negative stereotypes, one almost has to impose to decision makers additional interactions with the out-group. In a recent natural experiment, Shook and Fazio (2008) demonstrated how such additional interactions can reduce racial prejudice. They analyzed the evolution of the attitudes of White freshmen that were randomly assigned an African American roommate or a White roommates. The racial attitudes (toward African Americans) of those with an African American roommate became more positive after one quarter but the racial attitudes of those with a White roommate did not change. To the extent that prejudiced attitudes can be seen as an illusory correlation between group membership and the assessment of a positively valued attribute, this field study suggests that additional exposure can help correct illusory correlations.

6.4. Positive vs. negative illusory correlations

So far, we have framed our discussion of the emergence of illusory correlations mostly around cases where the illusory correlation is positive. But it is worth noting that a positive illusory correlation between two positively valued attributes can be seen as a negative illusory correlation between one positively valued attribute and one negatively valued attribute. For example, the positive correlation between proximity and attitude can be seen as a negative correlation between distance and attitude (as discussed in the previous section, Theorem 1 is still valid even if the sampling probability is not increasing in the attributes, provided that the log-submodularity or log-supermodularity conditions are satisfied).

But a far more interesting aspect of our analysis is that our model can also lead to the emergence of negative illusory correlations when both attributes are valued positively. As formalized in Proposition 2, negative illusory correlations are likely to emerge in settings where the attributes are considered as complements in the evaluation of an alternative. To illustrate this, consider an academic department with a hiring policy of only hiring faculty members who are both good teachers and productive researchers. Our model suggests that this will lead to the emergence of an illusory negative correlation: members of the department will come to underestimate the correlation between the two skills (or to believe it is negative if it is zero in reality). To explain why, note that the hiring committee will reject and most likely receive little further information about candidates that are not perceived as both good teachers and productive researchers. Candidates that are perceived to be good on both of these criteria will be hired. Over time, members of the hiring committee will find out whether the individuals hired are in fact good teachers and productive researchers; some will be found wanting. Overall this implies that the proportion of individuals believed to possess both skills will be underestimated relative to the true proportion. The proportion of individuals lacking at least one skill will, however, generally not be underestimated. All in all, this pattern corresponds to a negative illusory correlation.

Contrast this to what would happen with a department that is satisfied with candidates who are good on just one dimension. In such a setting, the two skills are seen as substitutes, and our model suggests that it will lead to the emergence of a positive illusory correlation between teaching competence and research productivity. In this case, candidates who are perceived to fall short on at least one dimension are rejected and probably not heard from again. As a result, estimates about their skill levels will remain stable, even if mistaken. Overall this implies that the proportions of individuals believed to lack at least one skill will be overestimated relative to the true proportion. The proportion of individuals having the two skills will, however, not be overestimated. All in all, this pattern corresponds to a positive illusory correlation.

A similar effect will emerge in other settings, such as when a manager is evaluating the quality of service providers on several attributes.

7. Conclusion

People who make sampling decisions on the basis of past experiences will typically obtain unrepresentative samples of information about the attributes of the choice alternatives. Those who seek positive experiences, and thus avoid alternatives that lead to poor outcomes, will tend to underestimate the prevalence of positively valued distal attributes. Besides, the correlation between estimates of distal attributes will generally fail to reflect the underlying true correlation. Depending on how decision-makers combine the attribute estimates in making their sampling decisions, positive or even negative illusory correlations will emerge.

A distinctive feature of our analysis is that it assumes that people are good processors of information. In fact, our model underscores the fact that illusory correlations might emerge even in the absence of information processing bias – even if the decision maker is a naïve intuitive statistician (Juslin, Winman, & Hansson, 2007). More generally, we have made only minimal assumptions about the learning mechanism and how information is stored in memory. Our main result holds when the attribute estimates are any weighted average of past observations of the attributes. This suggests that almost any reinforcement

learning algorithm will lead to the emergence of illusory correlations, provided that further observations of an attribute are contingent on the estimates of the other attributes.

Rather than challenging existing explanations that rely on information processing biases, our model provides a complementary perspective to explaining systematic judgment biases observed in the field, such as the tendency to have more positive impressions toward proximate others or the formation of some stereotypes. Standard explanations focus on how errors in how people process information in the available sample can produce illusory correlations. Our explanation focuses on why the available sample might be biased and how this can lead to illusory correlations even in the absence of information processing biases. In reality, both mechanisms likely operate together. Information processing biases can lead to misperceptions of correlations in a sample that is itself already biased as a result of adaptive sampling. More work is clearly needed to find out in which situations the sampling bias is the key. But where the sampling bias is the key factor, de-biasing how people process available information is just not enough to eliminate illusory correlations.

Acknowledgments

We thank Robin Hogarth for discussion and comments. This paper also benefited from comments from participants at the Management and Behavioral Research Breakfast at Universitat Pompeu Fabra, the Society for Personality and Social Psychology Conference, 2008, the Behavioral Decision Research in Management Conference, 2008, the 30th Annual Conference of the Cognitive Science Society and the 41st Annual Meeting of the Society for Mathematical Psychology, 2008. Gaël Le Mens was supported by a Stanford Graduate School of Business Fellowship and by a Juan de la Cierva Fellowship from the Spanish Ministry of Science and Education.

Appendix A. Proof of Proposition 2

We have:

$$\frac{\partial^2 \ln Q(x,y)}{\partial x \partial y} = -k[e^{r-s\gamma}(r+p-1)+p-1],$$

where $r = s(\frac{x^p+y^p}{2})^{\frac{1}{p}}$ and

$$k = \frac{rx^{p-1}y^{p-1}}{(1+e^{r-s\gamma})^2(x^p+y^p)^2} > 0.$$

Whenever $p \geqslant 1$, then $e^{r-s\gamma}(r+p-1)+p-1>0$ and the correlation is positive by Theorem 1 (i).

Next, we show there is a value of $p = p^* < 1$, such that $\partial^2 \ln Q(x,y)/\partial x \partial y > 0$ for all $p < p^*$. Note that $r \in [0, s\alpha]$. When p is low enough, with $p < 0$, $e^{r-s\gamma}(r+p-1)+p-1$ is negative because r is bounded. Hence, $\frac{\partial^2 \ln Q(x,y)}{\partial x \partial y}$ is positive for p low enough and the correlation is negative by Theorem 1 (ii).

References

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*, 27–58.

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin, 114*, 435–448.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Busemeyer, J. R., & Myung, I. J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General, 121*, 177–194.

Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. *Journal of Abnormal Psychology, 72*, 193–204.

Cooper, W. H. (1981). Ubiquitous halo. *Psychological Bulletin, 90*(2), 218–244.

Dawes, R. (1964). Social selection based on multidimensional criteria. *Journal of Abnormal and Social Psychology, 68*, 104–109.

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review, 112*, 951–978.

Denrell, J., & Le Mens, G. (2007). Interdependent sampling and social influence. *Psychological Review, 114*, 398–422.

Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision-making. *Psychological Bulletin, 73*, 221–230.

Fazio, R. H., Eiser, J. R., & Shook, N. J. (2004). Attitude formation through exploration: Valence asymmetries. *Journal of Personality and Social Psychology, 87*, 293–311.

Ferris, G. R., Judge, T. A., Rowland, K. M., & Fitzgibbons, D. E. (1994). Subordinate influence and the performance evaluation process: Test of a model. *Organizational Behavior and Human Decision Processes, 58*, 101–135.

Festinger, L., Schachter, S., & Back, K. W. (1950). *Social pressures in informal groups: A study of human factors in housing*. Stanford, CA: Stanford University Press.

Fiedler, K. (1991). The tricky nature of skewed frequency tables: An information loss account of distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology, 60*, 24–36.

Fiedler, K. (2000a). Illusory correlations: A simple associative algorithm provides a convergent account of seemingly divergent paradigms. *Review of General Psychology, 4*, 25–58.

Fiedler, K. (2000b). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review, 107*, 659–676.

Fiedler, K., Walther, E., & Nickel, S. (1999). The auto-verification of social hypotheses: Stereotyping and the power of sample size. *Journal of Personality & Social Psychology, 77*, 5–18.

Fiedler, K., Walther, E., Freytag, P., & Plessner, H. (2002). Judgment biases in a simulated classroom: A cognitive-environmental approach. *Organizational Behavior and Human Decision Processes, 88*, 527–561.

Fiedler, K., & Freytag, P. (2004). Pseudocontingencies. *Journal of Personality and Social Psychology, 87*, 453–467.

Fiedler, K., Freytag, P., & Unkelbach, C. (2007). Pseudocontingencies in a simulated classroom. *Journal of Personality and Social Psychology, 92*, 665–677.

Fiedler, K., Freytag, P., & Meiser, T. (2009). Pseudocontingencies: An integrative account of an intriguing cognitive illusion. *Psychological Review, 116*, 187–206.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12*, 392–407.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575–604.

Hogarth, R. M., & Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive Psychology, 24*, 1–55.

Juslin, P., Winman, A., & Hansson, P. (2007). The naive intuitive statistician: A naive sampling model of intuitive confidence intervals. *Psychological Review, 114*, 678.

Kashima, Y., & Kerekes, A. R. Z. (1994). A distributed memory model for averaging phenomena in person impression formation. *Journal of Experimental Social Psychology, 30*, 407–455.

Kareev, Y. (1995a). Positive bias in the perception of covariation. *Psychological Review, 102*, 490–502.

Kareev, Y. (1995b). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition, 56*, 263–269.

Kareev, Y. (2000). Seven (indeed, plus or minus two) and the detection of correlation. *Psychological Review, 107*, 397–402.

Kareev, Y., & Fiedler, K. (2006). Nonproportional sampling and the amplification of correlations. *Psychological Science, 17*, 715–720.

Karlin, S., & Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities I. Multivariate totally positive distributions. *Journal of Multivariate Analysis, 10*, 467–498.

Keeney, R., & Raiffa, H. (1993). *Decisions with multiple objectives*. New York, NY, USA: Cambridge University Press.

Landier, A., Nair, V. B., & Wulf, J. (2007). Trade-offs in staying close: Corporate decision making and geographic dispersion. *Review of Financial Studies, 22*, 1119–1148.

Landy, D., & Sigall, H. (1974). Beauty is talent: Task evaluation as a function of the performer's physical attractiveness. *Journal of Personality and Social Psychology, 29*, 299–304.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*, 75–98.

Levin, S., van Laar, C., & Sidanius, J. (2003). The effects of ingroup and outgroup friendships on ethnic attitudes in college: A longitudinal study. *Group Processes & Intergroup Relations, 6*, 76–92.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.

March, J. G. (1996). Learning to be risk averse. *Psychological Review, 103*, 309–319.

Marmaros, D., & Sacerdote, B. (2006). How do friendships form? *The Quarterly Journal of Economics, 121*, 79–119.

Newcomb, T. (1931). An experiment designed to test the validity of a rating technique. *Journal of Educational Psychology, 22*, 279–289.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology, 35*(4), 250–256.

Peeters, G. (1983). Relational and informational patterns in social cognition. In W. Doise & S. Moscovici (Eds.), *Current issues in European social psychology* (pp. 201–237). Cambridge, UK: Cambridge University Press.

Segal, M. W. (1974). Alphabet and attraction: An unobtrusive measure of the effect of propinquity in a field setting. *Journal of Personality and Social Psychology, 30*, 654–657.

Shook, N. J., & Fazio, R. H. (2008). Interracial roommate relationships: An experimental field test of the contact hypothesis. *Psychological Science, 19*(7), 717–723.

Smith, E. R., & Collins, E. C. (2009). Contextualizing person perception: Distributed social cognition. *Psychological Review, 116*, 343–364.

Sutton, R., & Barto, A. G. (1998). *Reinforcement learning*. Cambridge, MA: The MIT Press.

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*, 25–29.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207–232.